A SURVEY OF GAME THEORY MODELS ON PEACE AND WAR

Barry O'Neill

Departments of Political Science and Economics
and Centre for International and Strategic Studies
York University, Toronto

York Centre for International and Strategic Studies
Occasional Paper Number 9
March 1990

Prepared for R. Aumann and S. Hart, eds. *Handbook of Game Theory*.
New York: Elsevier

TABLE OF CONTENTS

Acknowledgements

This work was done while I was a visiting fellow at the Program on International Security and Arms Control, Yale, and the Center for International Security and Arms Control, Stanford. I would like to thank Jerry Bracken, Steve Brams, Bruce Bueno de Mesquita, Joanne Gowa, David Lalman, Fred Mosteller, Sanders MacLane, James Morrow, Barry Nalebuff, Michael Nicholson, James G. Taylor and John Tukey for their information and suggestions.

Introduction

Opinions on game theory's role in peace and war studies have been polarized from the first. McDonald (1949) praised the new discipline as `more avant-garde than Sartre, more subtle than a Jesuit,' and he quoted one Pentagon analyst, `We hope it will work, just as we hoped in 1942 that the atomic bomb would work.' P.M.S. Blackett (1961), who had organized British operational research during the war, was a skeptic, `I think the influence of game theory has been almost wholly detrimental.' If it had practical relevance, he claimed, it would have been snapped up by investors and card players, but it has not been, so `it is clearly useless for the much more complicated problems of war.'

This paper will present a survey of game theoretical applications to peace and war relevant to the continuing debate on the theory's place. (Some contributions are by Deutsch, 1954, 1968; Waltz, 1959; Quandt, 1961; Snyder, 1961; Shubik, 1968; Robinson, 1970; Rosenau, 1971; Junne, 1972; George and Smoke, 1974; Plon, 1976; Martin, 1978; Wagner, 1983; Maoz, 1985; Snidal, 1985a; Hardin, 1986; Larson, 1987; Jervis, 1988a; O'Neill, 1989b; and Rapoport, 1989.) The review will be non-evaluative, and will focus on the areas chosen for applications rather than developments in the mathematics. It will be fairly comprehensive in the international relations (IR) section, and include the main subjects in the military operations part. In regard to IR, I examine the mutual influences of the mathematics and the conventional theory or policy questions. The military section notes the interaction of game applications with new military strategy and technology. A companion paper (O'Neill, 1990b) surveys introductory writings for each game theory subfield that might be relevant to IR.

International Relations Applications of Game Theory

The Myth That Game Theory Shaped Nuclear Strategy

It has now become the common wisdom that in the late 1940s and 1950s game theory strongly influenced thinking on nuclear strategy. Nye and Lynn-Jones (1988) for example, write that as the field of strategic studies was developing around East-West problems, `Deterrence theory and game theory provided a powerful unifying framework for those central issues, but often at a cost of losing sight of the political and historical context' (p.6). The current view, they say, holds that `the abstract formulations of deterrence theory -- often derived from game theory - - disregard political realities' (p.11). Evangelista (1988, p.270) writes, `the early study of the postwar Soviet-American arms race was dominated by game theory and strategic rational-actor approaches.' Irving Louis Horowitz (1970, p.270) suggests that game theory became the `operational codebook' of the Kennedy administration. In his authoritative history of nuclear strategy, Freedman (1983, p.181) states, `for a time, until the mid-1960s, the employment of matrices was the *sine qua non* of a serious strategist.'

Statements like these suggest that the strategists used game theory's logic to derive their conclusions. In fact, examples of this are negligible in number. With a couple of exceptions, substantial game models of international strategy appeared only in the later 1960s after the tenets of nuclear strategy had developed. Like most writers who repeat this claim, Nye and Lynn-Jones give no supporting references. When Freedman asserts that until the mid-1960s any serious strategist used matrix games, can we believe that Brodie, Kahn, Kaufmann, Kissinger or Wohlstetter were less than serious? Hedley Bull, who associated with this group, wrote (1967, p.601), `the great majority of civilian strategists have made no use at all of game theory and indeed would be at a loss to give any account of it.' The historical record on game theory models supports his assessment. Freedman goes on to present the 2x2 Chicken matrix as a game used by early analysts. The irony is that not *until* the mid-1960s did the mathematical game of Chicken appear. Its sources were not nuclear planners, but two noted peace activists: Anatol Rapoport in 1965 defined `Chicken' as the now-familiar 2x2 matrix, Bertrand Russell in 1957 having used the teenage dare as an analogy for the arms race (O'Neill, 1988c). Russell's highway metaphor was prominent in the early discussions of strategists Schelling, Kahn and others, but the matrix game was not yet defined.

When one follows up citations purported to show game theory's importance, most fall into a few categories of non-applications. Some employ wargaming or systems analysis, rather than game theory. Some deal with military tactics, like submarine search or fire control, rather than strategy or foreign policy. Some use parlour games like monopoly, chess or poker as metaphors for international affairs, but give no formal analysis. Some only explain the principles of mathematical games and suggest that these might lead to a model (e.g., M. Kaplan, 1957). Others construct matrices, but analyse them only by decision theory. In particular, Ellsberg (1961) and G. Snyder (1961), writing on deterrence, and Schelling (1958a, 1960) on threat credibility, calculated no game-theoretical equilibria. A decision-maker in their models maximized an expectation against *postulated* probabilities of the adversary's moves. This mathematics had been available for centuries, and it was precisely to avoid assumptions about the other's move probabilities, that the founders of game theory developed minimax strategies and equilibria. In the game approach, the adversary's move probabilities are *deduced* from the strategic structure of the situation.

In their influential book on deterrence (1974) George and Smoke discuss game theory's contribution, and for their example, explain Snyder's 1961 model in a `modified version' (p.67). Comparing their account with the original, however, one sees that their modification was to drop Snyder's hypothetical probabilities of a Soviet attack, thereby converting his analysis from a decision to a game. While literally accurate, their treatment gives the impression that game theory modelling began earlier than it really did.

A final type of non-application involves strategists who approached their problem `with game theory principles in mind', by considering the adversary's perspective. As an influence of game theory, Fred Kaplan (1983) offers Albert Wohlstetter's judgement that American bomber bases near the Soviet Union were vulnerable to surprise attack. Of course, thinking from the enemy's viewpoint must be as old as warfare, and cannot be credited to von Neumann and Morgenstern. Kaplan makes a lengthy attempt to link nuclear strategy and game theory, but in the end his case is no more than suggestive, and ignores the problem of evidence that should be at hand but is not. He describes the enthusiasm of mathematician John Williams to assemble a team at RAND, but evidently these hopes never materialized, since almost no game models of international relations or grand strategy from RAND have come forth or been cited. The applications there addressed military tactics. The only arguable exceptions were Schelling's preemptive instability model (1958a) and the borderline model of Goldhamer *et al*., (1959), described later, but the latter left out information necessary for a well-defined game, and derived no optimal strategies.

As a criterion for what counts as a `game model', one should expect at least a precise statement of assumptions and a derivation of the parties' best strategies, either for prediction or illustration, but one encounters almost no such examples. The only publications of game models in the classical nuclear strategy literature were Morton Kaplan's deterrence model (1958) and Schelling's *Strategy of Conflict* (1960), which used 2x2 matrices to explicate trust and promises, and model first-strike instability. The book has been most influential, but the importance of these particular models should not be exaggerated: they were elementary and were untouched by other authors for many years. Schelling had two themes, the reorientation of game theory and military/strategic analysis, but he did not claim to derive the latter from the former. Their juxtaposition in his book had the unintentional but lasting effect of bolstering the misconception that nuclear strategy was somehow based on mathematics. This myth has been important in the history of strategic studies, but game theory itself has not.

The myth seems to have begun in the late 1940s with expositions like McDonald's, suggesting that the new theory would reveal how to exploit the new weapon. The activities of game theory's founders gave it credibility, Morgenstern publishing on defence matters and von Neumann advising the US government on nuclear weapons and ICBMs. The Hollywood movie WarGames depicted a mathematical genius who built giant computers to control the US missile force and wrote papers connecting poker and nuclear war. In fact von Neumann's thinking on international affairs avoided game concepts (York, 1987, p.89), and Morgenstern's political writings included parlour game analogies and mathematical-sounding terms, but no models (e.g., 1959). The 1960s debate on the validity of nuclear strategy reinforced the myth. The anti's attacked game theory and nuclear strategic thinking together, as acontextual, ahistorical, positivistic and counter to common sense. Supporters of nuclear strategy embraced deductive analysis, and this heading fit both game theory and nuclear strategy.

Neither group pinpointed just how game theory had influenced nuclear thinking. The most extensive critiques of strategic thought (Rapoport, 1964; Green, 1966) did not in fact claim that game theory was widely used to formulate it. Rapoport's thesis was more subtle; he claimed that the very existence of a formal theory of conflict contributed to the expectation that nuclear dilemmas could be solved by *a priori* reasoning. However, many asserted a real link without specifying just how game theory was involved (e.g., Zuckermann, 1961; Horowitz, 1962, 1963a,b, 1970).

Indeed, there was strong evidence available against a link. In a survey of 177 non-governmental nuclear strategists, Licklider (1971) found only two mathematicians. Donald Brennan (1965), a prominent strategist at the Hudson Institute, estimated that the reports issued by his centre, classified and unclassified, might fill twenty feet of shelf space, but `the number of individual pages on which there is any discussion of concepts from game theory could be counted on one hand; there has been no use at all made of game theory in a formal or quantitative sense. . . . "gaming" is much used by nuclear strategists, "game theory" is not.'

The notion that game theory has already been tried and left behind, probably dampens current interest. We should realize that the many models of the last few years are a first flowering. The most appropriate mathematical techniques, like incomplete information or equilibrium selection in dynamic games, were invented fairly recently. They have initiated this expansion, and we can be optimistic about future findings.

The Debate On International Cooperation And Realism

The dominant philosophy in Western IR theory is `realism' (Keohane, 1986a). The actors are taken to be states, who pursue their national interests in an anarchic world. To survive the constant peril of subjugation, they seek power, which in a competitive system is defined relative to other states, and so can never be achieved mutually and finally. Significant cooperation is impossible, or, for `neorealists,' very difficult. Alternatives to realism hold a greater expectation of cooperation, either because of the growth of international institutions and interdependence (traditional idealism), the imposition of order by a dominant power (hegemonic stability theory), the common interests of subregions (regional integration theory), standards of international behaviour (regime theory), or transnational class interests (Marxist structuralism).

Prisoner's Dilemma (PD) games have formalized a central issue between realists and regime theorists: How can cooperation develop in an anarchic world? Keohane (1985, 1986b) explicates his regime theory position with simple 2x2 games, and Taylor's work (1976, 1987), although not addressed specifically to IR, has been influential. Axelrod (1984) discussed cooperation in repeated PD with probabilistic termination, comparing it with war and competition between species. The IR theory aspects of finite or repeated PD have been treated informally by Axelrod and Keohane (1985), Brams (1985), Grieco (1988), Jervis (1988), Larson (1987), Snidal (1985a,b), and Weber (1988), and the authors in Oye's volume (1986). Stein (1982a) compares various matrix representations, Snidal (1985b, 1989) uses n-person and non-PD games to analyse international cooperation, and McGinnis (1986) considers issue linkage, a central idea of regime theory. It is this area of the growth of cooperation that has seen the most interaction of mathematical and non-mathematical researchers, and produced papers of many different degrees of formalization. The repeated PD model was interesting and accessible and appeared when the political theory question was prominent.

Relevant to the realist debate on cooperation are game models in international political economy, in the sense of research on international trade and exchange. Some cleverly interweave empirical details and conventional theory with simple matrix games. Conybeare's work (1984, 1987) separates trade wars into Prisoner's Dilemma versus Chicken types. Alt, Calvert and Humes (1988), Gowa (1986, 1989) and Snidal (1985c) use matrix games extensively in their analyses of hegemony. Wagner (1988a) uses Nash bargaining theory to analyse how economic interdependence can be converted to political influence. Kivikara and Nurmi (1986) and Baldwin and Clarke (1987) model trade negotiations. A frequent theme has been the Oil

Producing Exporting Countries (Bird and Sampson, 1976; Brown, Mearsheimer and Petersen, 1976; Bird, 1977; Sampson, 1982), often analyzed as a coalitional game. One neglected area is the use of game-theoretical analyses of voting to propose decision rules for international organizations, the lone example being Weber and Wiesmeth's paper (1989a).

A central realist concept is that of power in the international system, and one model is Selten and Mayberry's (1968). They postulate two countries who build weapons at linear costs, and share a prize according to the fraction that each holds of the total military force. The prize might involve political compliance from the adversary or from third countries. Brito and Intriligator (1977) have nations invest money in weapons, to increase their share of the international pie, as determined by a version of Nash bargaining theory. These arsenals increase the risk of war. Shapley and Aumann (1977) revise standard bargaining theory so that threats involve a non-refundable cost, e.g., the money spent on arms. Brito and Intriligator (1985) postulate countries that commit themselves to a retaliatory strategy to be implemented in the second period in order to induce a redistribution of goods in the first. Their model is also applied by Don (1983). Matsubara (1989) uses the same structure for a theoretical discussion of international power. Zellerman (1961) and Gillespie and Zinnes (1976, 1977) discuss large countries' attempts to gain influence in the Third World, the latter authors applying differential games. Mares (1988) discusses how a large power may compel a middle one, examining the US versus Argentina during the Second World War and later Mexico. Spandau (1978) gives a repeated PD model of an economic boycott of South Africa. Bueno de Mesquita (1990) and Bueno de Mesquita and Lalman (1990a) examine how wars change the hegemony pattern in the international system, exemplified by the Seven Weeks War of 1866.

According to the realist approach, one can explain the gist of international relations by taking states as the dominant actors and unitary ones. Achen (1988) gives conditions on a state's decision-making process such that it would act as if it had a utility function, even though its behaviour is being set by competing domestic groups.

Many in the IR mainstream view game theory as naturally realist-oriented in that it construes governments as the players and their choices as rational. Rosenau (1971), for one, writes, `The external behaviour called for in game-theoretical models, for example, presumes rational decision-makers who are impervious to the need to placate their domestic opponents or, indeed, to any influences other than the strategic requisites of responding to adversaries abroad.' On the contrary, one can cite studies that emphasize domestic factions, or portray governments as acting against their sober interests. Brito and Intriligator (1980) and Intriligator and Brito (1984) suggest a coalitional game with six players: the governments, electorates and military-industrial complexes of the United States and the Soviet Union. Although the two military-industrial complexes cannot directly collude, they know how to coordinate to forward their common goals. In 1968, Eisner gave a timely treatment of Vietnam War policy, with North Vietnam, President Johnson, and the American public as players. Bueno de Mesquita and Lalman's (1990a) investigate the influence of domestic politics on crisis bargaining. Sandretto's players (1975) include multinational corporations. He extends the theories of Marx, Lenin and Rosa Luxemburg on how transnational control of the means of production influences world politics.

There is an irony in the common belief that game theory requires confidence in everyone's `rationality': the very first significant game model, Schelling's analysis of pre-emptive instability (1958b, 1960), had each side worrying that the other would attack irrationally. O'Neill's players in a model of nuclear escalation in NATO (1990c) act from transitory emotions of anger and fear. These and other counter-examples show that it is not game theory *per se* that determines whether `national interests' or any other motive are the players' goals, or whether states or non-state entities constitute the players. The choice is at the modeller's discretion. The essence of game theory is rather that each player considers the other's view of the conflict as part of its own choice of strategy.

Negotiations On Arms Control And Other Issues

The first game applications in this area were mathematically auspicious, but too abstract to influence mainstream political analysis. At the instigation of director Herbert Scoville, the United States Arms Control and Disarmament Agency funded many of the prominent theorists to investigate arms control bargaining, and these papers were the founding studies on repeated games of incomplete information (Aumann *et al*., 1966, 1967, 1968). Saaty (1968) summarizes one of their models. Their work was too technical to influence political science modelling, and the approach has been taken up again only recently with the work of Downs and Rocke (1987, 1990), who study arms control negotiations using repeated Prisoner's Dilemma.

Less formally, Weber (1988) uses Axelrod's approach to examine restraint on deploying anti-missile missiles, anti-satellite weapons and multiple warhead ICBMs. Picard (1979), Makins (1984) and Bunn and Payne (1988) study the strategic arms negotiations as simple games, and Jonsson (1976) the nuclear test ban. Hopmann (1978) uses Harsanyi's model of power to conventional arms talks. Brams (1989) and Guth (1985, 1988) treat the NATO issue that negotiating with a rival requires negotiating concurrently with one's allies.

No studies have compared game models of bargaining with real international negotiators, probably because until recently the mathematical theory has emphasized the outcome and trivialized the dynamics of negotiation. Several authors have suggested clever mechanisms to reach an accord. Kettelle (1985) proposes that each side feed its best alternative to agreement into a computer that simply announces whether a compromise is possible if negotiations continue. A `go ahead' might prompt more effort to find a compromise. Unaware of the theoretical literature, he did not ask whether each side had an incentive be honest with the computer. Green and Laffont (1988; summarized by O'Neill, 1990d) outline a procedure in which the parties have only one agreement available, to take or leave. In a series of rounds each says yes or no, until both accept or time runs out. Each is not sure whether the agreement will help it or hurt it, but has partial knowledge relevant to its own and the other's gains, and at the equilibrium strategies, can infer the other's knowledge from the other's delays in accepting. Frey (1978) discusses an `insurance' system to convert PD traps of arms-building into efficient ones prompting agreements to reduce arms. Vickrey (1978) applies the idea of the Clarke tax to induce governments to state their true interests, and Isard (1965) presents a technique based on metagames. Singer (1963) and Salter (1986) expand an application originally proposed by Kuhn: each side separates its arms into piles, and the other chooses which pile must be dismantled. Like divide-and-choose, the goal is a situation where a side that perceives its military position to have suffered, can only blame itself, but what assumptions about how weapons interact in war would allow such a complaint-proof disarmament process? The idea is ripe for formal development.

Models Of Arms Building

An influential analysis is Jervis' explication (1978) of the security dilemma, in which each side tries to increase its own security, and thereby lowers the other's. If technology could develop a purely defensive weapon, one that had no role in offensive campaigns, the dilemma would disappear, but typically weapons are useful for both protection and aggression. Jervis modelled the security dilemma by the 2x2 game shown below, where the two moves might correspond to passing up some new military system or building it, `4' means `most preferred', and the two equilibria are underlined.

$\underline{4,4}$	1,3
3,1	$\underline{2,2}$

His paper drew attention to this game, usually called the Stag Hunt, sometimes Assurance or Reciprocity, which had been overshadowed by Prisoner's Dilemma and Chicken. It is interesting that political scientists recognized its importance before game theorists. Most game theorists had never heard the name or dismissed it as trivial because one equilibrium is ideal for both players, but some recent theoretical papers have taken it seriously following Harsanyi and Selten's (1988) and Aumann's (1987) argument that the upper left outcome may not be self-enforcing, even though it is the payoff-dominating equilibrium. Apparently governments do end up instead at the poor equilibrium, each fearing the other will choose the second strategy.

Political scientists extending Jervis' work are Sharp (1986) and Snyder (1984). Avenhaus *et al*. (1988) develop the security dilemma, giving each government the choice of building offence- or defence-oriented weapons, which the other will observe to judge aggressive intentions. Guth and van Damme (1989) analyse how a choice of force structure can signal one's intent. O'Neill's model (1988b) relates the degree of offensive advantage to the benefits of military secrecy, another theme in the security dilemma literature.

Lewis Frye Richardson's differential equations of arms races have fascinated so many scholars, they were sure to appear in a differential game (Deyer and Sen, 1984; Gillespie and Zinnes, 1977; Simaan and Cruz, 1973, 1975a,b, 1977; Zinnes *et al*., 1978; see also Case, 1979, Moriarty, 1984, and Saaty, 1968). Melese and Michel (1989) give a version with a closed form solution. Another substantial literature comprises repeated PD models of arms races (Brams, 1985; Downs, *et al*., 1985; Lichbach, 1988a,b; Majeski, 1984, 1986a,b; McGinnis, 1986, 1988a,b; Stubbs, 1988). In a series of models, many collected in their book, Brams and Kilgour (1987b, 1988, 1990) ask how threats of retaliation can support a cooperative outcome. They investigate multistage games involving plays of continuous versions of a Chicken or Prisoner's Dilemma matrix, where players adopt degrees of cooperation, followed by possible retaliation.

A fundamental question is whether to keep one's arsenal a secret (Brams, 1985). One might want to exaggerate it to deter an adversary, or understate it to avoid provoking the other or gain a surprise advantage. Putting the question positively, will the new Soviet openness to verification allow the superpowers to cut most of their strategic weapons? In Nalebuff's model (1986) secrecy is harmful, and in Sobel's (1988) keeping secrets would help but is not the equilibrium move. At the sensible equilibrium the two sides reveal their holdings, even though both sides would gain by mutual secrecy.

Sobel assumes two aggressive governments, each possessing a military strength that is known to itself and can be verifiably revealed to the other. Each decides whether to reveal, and then chooses either to attack or to adopt a defensive posture. If there is a war, whoever has the larger force wins, and losing is less harmful if the loser has adopted a defensive posture. The payoffs to X decrease in this order: 1) X and Y attack, X wins; 2) X attacks and Y defends, X wins; 3) both defend, so there is no war; 4) X defends and Y attacks, X loses; and 5) X attacks and Y attacks, X loses. The payoffs to Y are analogous. Sobel's pessimistic argument looks for a non-extreme cutoff such that those of military strength greater than the cutoff will reveal it, and those lower will not. However an equilibrium pair of such cutoffs cannot exist, because a government just slightly below the cutoff will not want its strength to be underestimated as that of the typical government below. This underestimate would increase its risk of being attacked. It will reveal its strength, and thus the set of those practising secrecy will `leak from the top end.' For some parameter values this is an unfortunate outcome, since within the model, knowledge of the other's power leads the stronger to attack.

I will mention only examples of the many papers using experiments or simulations on game matrices to clarify the arms race. Alker and Hurwitz paper (n.d.) discusses PD experiments; Homer-Dixon and Oliveau's After-MAD game (Dewdney, 1987) is primarily educational; Alker *et al*. (1980) analyse a computer simulation on international cooperation based on repeated PD; experiments like those of Pilisuk (1984), Pilisuk and Rapoport (1963) and Plous (1987) interpret game moves as arming and disarming, escalating or de-escalating; and Lindskold's experiment (1978) studies GRIT, the tit-for-tat tension-reducing procedure.

Deterrence

The notion of deterrence has generated hundreds of political science articles, a phenomenon that might puzzle a game theorist. Their focus is credibility, since many nuclear threats are not worth carrying out. One example would be NATO's policy of initiating the use of nuclear weapons if it is losing a conventional war in Europe. However, in the most straightforward game models, credibility is trivialized as all or none depending on the payoffs: a threat will dissuade the adversary if and only if going through with it is in one's interests. The concept `deterrence' is generally not used in game theory, at least not in connection with retaliatory threats, nor are the associated notions of `credibility' and `resolve.' Instead the focus has been on `reputation,' the other's belief in the unlikely but possible circumstance that the threatener actually prefers retaliating to not, and the course of this belief as the game is repeated. However, the assumptions of repeated plays and evolving reputations do not fit international crises, where issues differ from one to the next, government officials come and go, and in the case of the ultimate threat, the game can not be repeated.

Security scholars have taken other approaches, following especially the ideas of Schelling (1960, 1966, 1967). The key problem of nuclear deterrence is the adversary's scepticism that one will carry out a suicidal threat. Schelling suggested two answers, alternatives to brandishing total destruction. The first was performing a series of small escalations that gradually increase the pain suffered by both antagonists. In United States nuclear strategy this approach has produced attempts to set up `limited nuclear options,' and deploy weapons that fill in `gaps in the escalation ladder.' Schelling's second method involves setting up a mechanism that causes no immediate pain but repeatedly generates a risk of total disaster, allowing the two sides to match each other in risk-taking until one gives in, like two people roped together moving down a slippery slope ending in a precipice. When a threat to commit murder-and-suicide would be ignored, a threat to incur a chance of disaster might work. It is doubtful that these schemes could really be implemented, but they have become the paradigms in nuclear deterrence theory.

Powell (1987a,b, 1989a,b,c, 1990) makes Schelling's ideas mathematically rigorous, using sequential games of incomplete information. A government's resolve may be high or low, and if the government stays in the conflict, the other will raise its estimate of this resolve. The term `resolve' is used vaguely in the mainstream literature, but Powell's models justify its definition as a ratio of certain utility increments, since this ratio is a determinant of who will prevail. He discusses how the conflict outcome is influenced by perceptions of resolve, not just actual resolve, how the information held by the parties affects the likelihood of inadvertent war, and how a change in the size of the escalatory steps alters the likelihood of war. He translates his solutions back into the vocabulary of security studies.

The common concept of `showing' resolve is full of puzzles: Clearly the weaker you are the more important it is to show resolve, so why are aggressive actions not interpreted as signs of weakness (Jervis, 1988b)? In many cases, weak and strong alike would make a display of resolve, perhaps because the action costs little, and because no action would be taken surely to mean weakness. But what rational basis is there for inferring weakness from inaction, especially if the weak always act? Inferences from moves no one would make are the domain of equilibrium refinement theory, and Nalebuff (1990) applies it in a simple model of promoting reputation. A situation presents you with the choice of acting (A) or not acting (N), and acting confers on you a benefit of x, known to all, and a cost of c, known only to you. Letting $\underline{c} = \underline{c}_A$ or $\underline{c}_N$ be everyone else's estimate of your cost following their observation that you did or did not act, you will receive a `reputation for toughness' payoff of $a(1-\underline{c})$, where $a>0$. Thus, acting yields you $x-c+a(1-\underline{c}_A)$ and not acting, $a(1-\underline{c}_N)$.

Nalebuff's structure is not really a game -- one side simply observes -- but it illustrates how different equilibrium theories restrict the others' reasonable beliefs about your cost. Assume that $x=1/4$, $a=1/2$, and that others hold a uniform distribution for c before they observe your move. One sequential equilibrium (Kreps and Wilson, 1982) calls for A if $c < 1/2$, N otherwise, so those who can act cheaply will do so. In another, you should act no matter what your type, since it is assumed others are disposed to believe a non-actor has the maximum cost of 1. A further sequential equilibrium calls for N for all types, since if you were to act, it is assumed that would believe that your cost is 1, that your eagerness to prove yourself betrays weakness.

Refinement theories eliminate some of these equilibria. Pearce's (1982) and Bernheim's (1982) rationalizability condition involves sequential elimination of dominated strategies, and this, it turns out, rejects the equilibrium of not acting for any cost. For the parameters of the example, the most that acting could benefit you through reputation is 1/2, so if others see you act and know you are sensible, they should decide that your cost must be no higher than 1/2, certainly not 1. A cost of 1/2 is the same cost as they would estimate for you if you did not act, so even when other's beliefs are worst for you, acting does not change your reputation, and you should act only if it is worth it disregarding reputation, i.e., if $c < 1/4$.

Nalebuff applies the arguments of universally divine equilibria (Banks and Sobel, 1987) and perfect sequential equilibria (Grossman and Perry, 1987) to eliminate further solutions, the latter refinement striking out the policy of always acting no matter what your cost. His approach adds to our understanding of both refinement theory and deterrence. By looking at a specific application, some government's attempt to signal, one can ask whether the refinement is reasonable in admitting some beliefs and rejecting others, and we are also led to seek out historical evidence on how a message of resolve was actually received. The latter question is neglected in the political science literature, and the few who have examined it, e.g., Thies (1980) on US signalling during the Vietnam war, have not supported the elaborate schemes of military buildups and attacks that governments have used to show resolve.

O'Neill's paper (1988b) complements Nalebuff's by discussing some non-informational rationales that might justify a display of resolve, including the volatile `war of face,' where the sides use amounts of their prestige as `bids' for a prize, and only the loser must pay. Brito and Intriligator (1987) present a further signalling model, and Banks (1990) gives a general theorem on when greater resolve actually leads to success in a crisis. O'Neill (1990f) suggests a mechanism behind large powers' preoccupation with commitment over issues of no innate importance, based on an analysis of the chivalric custom of challenging in the Middle Ages.

Wagner (1988b) and Zagare (1990) use game models and sequential equilibria to answer those political scientists who have criticized deterrence theory from psychological and historical perspectives. Bueno de Mesquita and Lalman (1989) discuss how each side's incomplete information about the other's intentions can lead to a war desired by neither, and argue that the confirmation of deterrence theory may be in part artifactual. Quester (1982) used these structures to categorize ways that wars can start, using historical examples. Stein (1980, 1982) looked at various incidents from the viewpoint of conditional rewards and punishments, and in the latter paper, at types of misperception. Wilson (1986) gives a repeated game model of strikes and counterstrikes in which a sequential equilibrium directs each side to maintain a positive probability of striking at each trial. Zagare (1987b) axiomatizes the structure of mutual deterrence situations and derives that it is a PD game, but Brams and Kilgour's (1988) critique of his assumptions lead them to Chicken.

To explicate partial credibility of a deterrent threat, O'Neill (1990c) posits a threatener who, when crossed, retaliates with a probability increasing with the damage done when the threat was ignored, and decreasing with the cost that the threatener would suffer from retaliating. The model is used to analyse justifications for NATO's intermediate-range nuclear missiles, which were intended to promote `coupling,' the tendency of a conventional war in Europe to trigger a global nuclear war, a goal of NATO's controversial strategy.

Wagner (1982) and Allan (1980, 1983) have formalized the idea that deterrence involves negotiation over the boundaries of power, over what rivals may do without facing retaliation. Examining multi-sided threats, Scarf (1967) assumes that each government allocates some of its weapons specifically against each of the others, and that each configuration of all parties' allocations yields a vector of utilities, measured ordinally. Adapting his n-person solution theory (1971), he states assumptions guaranteeing a core allocation of military threats. His theory is developed by Kalman (1971) and modified to a sequential game by Selten and Tietz (1972).

O'Neill (1989a) explicates Kull's `perception theory' (1988) of the superpowers' motives for building far more nuclear weapons than needed for deterrence, and choosing weapons suitable only for unrealistic counterforce wars. Each side knows that nuclear superiority and warfighting ability are militarily irrelevant, and knows that the adversary trusts this, but they are unsure that the adversary knows their own good sense in this regard. Using Aumann's approach (1988), which arose from the common knowledge literature, O'Neill finds a family of belief structures that induce them to build militarily useless armaments.

Balas (1978) calculates the appropriate size of the US Strategic Petroleum Reserve, proposed as a way to reduce vulnerability to oil embargoes. Comparing his analysis with a non-game-theoretical model based on an exogenously given likelihood of an embargo, one finds the latter approach tends to overestimate the optimal reserve size, because it neglects the greater deterrent effect due to greater size. Several papers investigate deterring and negotiating with international terrorists (e.g., Sandler and Lapan 1988).

Intriligator (1967) and Brito (1972) modelled a missile war as a differential game, with bounds on the rates of fire and the attrition of missiles and economic targets governed by differential equations. Intriligator's approach was developed by Chaatopadhyay (1969), and applied by himself, Brito and others in many subsequent articles, to derive a lower limit on armaments that guarantee deterrence.

Two French scholars, Bourqui (1984) and Rudnianski (1985, 1986), have investigated a concern of their country, how a small nuclear power can deter a large one. The notions of threats and deterrence in 2x2 games are treated also by Dacey (1987), Maoz and Fiesenthal (1987), Moulin (1981), Reisinger (1989), Snyder (1971) and Stefanek (1985). More general discussions of game theory and deterrence are by Kugler and Zagare (1987), P. Bracken (1984), and P. Bracken *et al*., (1984), and Gekker (1989b). Some matrix games portray the ethical dilemmas of nuclear weapons (Hardin, 1983, 1986a; Woodward, 1989). O'Neill (1989b) surveys the goals of game models of deterrence.

Crisis Instability And The Outbreak Of War

Schelling's preemptive instability model (1960) was apparently the first published non-trivial game application to international strategy. Anticipating the spirit of games of incomplete information, he assumed that each side feared that the other would attack irrationally, and traced the payoff-dominating equilibrium of a 2x2 game as this fear grew. Nicholson (1970) extended his model.

One way game theory has contributed to less formal disciplines is by sharpening existing concepts, sometimes by suggesting a numerical measure. Examples of measures are Harsanyi's definition of power (1962), Axelrod's measure for conflict of interest (1969), O'Neill's measure of advantage in a brinkmanship crisis (1988c), Powell's formula for resolve (1987a,b, 1990) and measures for the value of a weapon, discussed below. O'Neill offers a formula for degree of crisis instability (1987), the temptation for each government to attack due to fear that the other is about to, fuelled by a military advantage from striking first. The decision is cast as a Stag Hunt, and the measure is applied to the debate on space-based missile defences. Bracken (1989b) proposes an alternative measure for this class of games. Recent interest in missile defences in space has generated about three dozen quantitative papers asking whether these systems increase the temptation to strike (Oerlich and Bracken, 1988). In many cases their authors were grappling with a basically game-theoretical problem, but were not ready to apply the right technique, and so built fatal inconsistencies into the models.

In Nalebuff's model of crisis instability (1989), government A holds a function that states its probability that B will strike is a function of B's assessed likelihood that A will strike. Some exogenous crisis event impels these beliefs up to an equilibrium. Powell (1989a) states assumptions on a game of escalation with complete information such that there will be no temptation to strike, and his model is extended by Bracken and Shubik (1990). A related analysis is Franck's (1983). Bracken and Shubik (1989) look at the war consequences of all possible coalitions of the nuclear powers, in regard to preemptive stability.

Greenberg (1982) represents the decision of how to launch a surprise attack. O'Neill (1989c) models the `when' aspect of surprise along with the defender's decision to mobilize in the face of uncertain warning, using the stochastic processes literature on quickest detection. The goal is to produce a measure of the danger of a war through false alarm as a function of the parameter probabilities of the warning systems.

Escalation

Two views of escalation regard it as either a tool that each side manipulates to promote its goals, or a self-propelling force that takes control of the players. Models with complete information and foresight tend to the former picture, but those that place limits on the governments' knowledge and judgement usually predict that they will be swept up into the spiral of moves and responses. An example of the former is the dollar auction, a simple bidding game somewhat like international escalation in that both the winner and the loser must forfeit their bids. O'Neill (1986) and Leininger (1988, 1989) give the solution, in which one bids a certain amount determined by the resources held by each, and the other resigns. The rule is unintuitive and at variance with the experimental finding that people do bid up and up, but it exemplifies one role of game theory, to establish the baseline of what constitutes rational behaviour, in order to identify the deviation to be explained.

Langlois (1988, 1989) develops a repeated continuous game model postulating limited foresight in escalation: each side acts not move by move, but chooses a reaction function, a rule for how strongly it will respond to the other side's aggressiveness. He derives conditions for an equilibrium pair of functions, investigates when a crisis would escalate or calm down, and, in general, argues strongly for the repeated game approach. O'Neill (1988d) models limited foresight using the artificial intelligence theory of game-playing heuristics, to show that under some conditions players do worse by looking further down the tree.

Uscher (1987) compares repeated PD and Chicken models of escalation. Powell (1987a,b, 1990) analyses crisis bargaining in the context of his escalatory model, and Leng (1988) uses the theory of moves to discuss the learning of bargaining stances from the experience of past crises. Further studies on escalation are by Brown, *et al*. (1973) and Ackoff *et al*. (1968).

Morrow (1990) considers bargaining and escalating as complementary moves during a crisis, and has each side learn about the other's capacity for war through its delay of an agreement, or through its willingness to fight a small conflict. In a follow-on paper (1989a), he asks whether each state can learn enough about the adversary to avoid the next crisis. A block to resolving a crisis is that any side initiating a compromise may be seen as weak. Morrow (1989b) looks for information conditions that alleviate this problem, which is the diplomatic analogue of a military first-strike advantage.

Alliances

Arguing from coalition-form game theory, Riker (1962) suggested that a smallest winning coalition will be the one that forms. He discussed the implications for national and international affairs. In the theory of national politics his ideas were very influential, but were followed up only sparingly in the international context (Siverson and McCarty, 1978; Diskin and Mishal, 1984).

How will nations respond to an accumulation of power by one party? Will they `bandwagon,' i.e., join the powerful country to be on the winning side, or will they `balance,' i.e., form a countervailing alliance against the threatening power? Much of US foreign policy has followed the notion that small countries will bandwagon. Chatterjee (1975), Don (1986), Guner (1988), Maoz (1989a,b), Nicholson (1988), Luterbacher (1983), Hovi (1984), Selten (1988) and Niou and Ordeshook (1986, 1987, 1989a,b) have addressed this and related alliance questions using coalitional games, the latter two authors stating a solution theory developed from the nucleolus and bargaining set. In their book Niou, Ordeshook and Rose take pains to test the theory using large European power relations from 1871 until 1914.

Wagner (1987) addresses the traditional question of the optimal number of players to produce stability. Zinnes, Gillespie and Tahim (1978) give a differential game model of n-nation stability. Kupchan (1988) represents relations within an alliance by several 2x2 matrix games, while Sharp (1986) and Maoz (1988) discuss the dilemma between the greater security through an alliance versus greater danger of entanglement in a partner's wars. O'Neill (1990c) uses games of costly signalling to clarify the idea of `reassurance' in alliances. Allies invest in militarily pointless arms to show each other that they are motivated to lend support in a crisis.

The game literature on alliances has some significant gaps. Only one paper (Weber and Wiesmeth, 1989b) has used public goods game theory to treat the contentious question of alliance burdensharing. The new formal literature on coalition formation exemplified in Roth's volume (1988), has not been applied to military alliances, nor to `peaceful alliances,' like the signators to the Partial Test Ban Treaty or the nuclear Non-Proliferation Treaty.

Arms Control Verification

Game-theoretical studies of verification form two groups. The first and earlier involves decisions about allocating inspection resources or a quota of inspections limited by treaty, to catch an intelligent evader. The second is more abstract, treating general principles of whether to cheat and whether to accuse in the face of uncertain evidence.

Studies of how to allocate inspections began in the early 1960s. The nuclear test ban negotiations stalled on the issue of seven yearly inspections, the American demand, versus three, the Soviet offer, and a complete ban was never achieved. Maximizing the deterrent value of each inspection should please a government worried either about compliance or espionage, and this goal motivated the work of Davis (1963), Kuhn (1963), Dresher (1962) and Maschler (1966, 1967). Their inspector confronts a series of suspicious events, and each time faces the dilemma of whether to expend a visit from a quota, knowing that there would be fewer left thereafter. The problem is somewhat like a search game extended over time rather than space. Rapoport (1966) gives a simple example. Maschler's papers showed the advantage to an inspector of the ability to make a public commitment to a strategy before the evidence is examined. Moglewer (1973), Hopfinger (1975), Brams, Davis and Kilgour (1988) and Kilgour (1988) extend quota verification theory, and Filar (1983, 1984) and Filar and Schultz (1983) treat the case of an inspector who has different travel costs between different sites, these being known to the evader, who plans accordingly. Some recent sophisticated work has examined the monitoring of materials in nuclear energy plants, in support of the Non-Proliferation Treaty (Avenhaus, 1977, 1986; Avenhaus and Canty, 1987; Avenhaus, Fichtner and Vachon, 1987; Bierlein, 1983; Zamir, 1987.)

In the second group of more conceptual verification studies, typically the game is non-zerosum, and the decision is whether, rather than how, to violate and whether to accuse. Maschler's early paper (1963) is an example, and Bellany (1982), a political analyst, uses simple matrices to clarify the compliance decision. The evidence is sometimes represented as a single indicator that alleges `innocent' or `guilty' (Avenhaus and Frick, 1983; Fichtner, 1985, 1986; Dacey, 1979; Wittman, 1989), sometimes as a continuum of strength (O'Neill, 1990). Brams and his colleagues have developed several models (Brams, 1985; Brams and Davis, 1987; Brams and Kilgour, 1986, 1987a; Brams and Kilgour, 1987, 1988). O'Neill proves the odd result that with a more thorough verification scheme, the inspector may be less certain about the other's guilt in the light of guilty evidence, the reason being that with better verification, one's prior likelihood that the other would take the chance of violating, is lower.

Analyses Of Specific Situations

These studies usually apply the theory `straight from the book', without further modelling tricks specific to the application. By setting up a model of a real event, one cannot claim credit for prediction since the outcome is already on the record, but the method gives a narrative structure that organizes the account of each side's decisions.

The favourite subjects have been the Cuban Missile crisis, the US/Soviet arms confrontation, and the Middle East wars. A frequent approach is to portray the event as a small game, often a 2x2 matrix, and base the analysis only on the ordinal properties of the payoffs. Most authors use matrices from Rapoport and Guyer's taxonomy of 78 ordinal types (1966). Snyder and Deising (1977) did extensive and influential work on sixteen historical situations, using two by two matrices. Examples are Measor (1983), Boulding (1962) and Nicholson (1971) on the arms race, Gigandet (1983) on German/Soviet relations between the World Wars, Ravid (1981) on the decision to mobilize in the 1973 war, Forward (1971) and Weres (1982) on the Cuban missile crisis, Levy (1985) on the war in Namibia, Satwedin (1984) on the Thai-American relations in the early 1960s, and Kirby (1988) on anti-missile systems. Scholars from South Korea and West Germany have used matrix games to discuss problems of reunifying their respective countries (Pak, 1977; Gold, 1973; Krelle, 1975; see also Cook, 1975).

Some researchers have taken the trouble to interview subjects and apply scaling techniques in order to set the payoff entries. Delver (1986, 1987, 1989) contacted senior naval officers at the Royal Netherlands Naval Academy, and applied Dutch developments in bargaining theory to analyze the effect of NATO maritime strategy on superpower relations in Europe. Lumsden (1973), on the Cyprus crisis, interviewed Greek students, but was unable to talk to their Turkish counterparts. Plous (1985, 1988, 1989) sent questionnaires to all members of the US Senate, the Israeli Knesset, the Australian Parliament and the Politburo, asking their estimates of the benefits in case each superpower significantly reduced its nuclear arsenal. The Politburo was unresponsive in those pre-*glasnost* days, but thirty-two senators sent back data that ought to find use in our classroom examples.

Two-person matrices are accessible to the non-mathematical, but often they are just too simple to capture the structure of the conflict. In Plous' questionnaire, for example, it is implausible that one side would be moving with no information about what the other did. Snyder and Diesing depict many of their games as matrices, but the text becomes more understandable if one reinterprets the moves as sequential. Many applications should perhaps have selected games in extensive form or with incomplete information. A example of an extensive form analysis is Sexton and Young's (1985), on the Falklands/Malvinas War. Young (1983) argues strongly that this technique is no more complex but more adequate than 2x2 matrices. Johr (1971), a Swiss political scientist, uses extensive-form games in an historically intricate analysis of his country's relations with Nazi Germany. Zagare (1977) models the Vietnam war negotiations with a 3x3x3 game. Examples of incomplete information analyses of specific situations are Guth (1986) on negotiations over an intermediate missile ban in Europe, Wagner (1989) on the Cuban missile crisis, and Mishal *et al*. (1987) on the PLO/Israel conflict. Hashim (1984), at the University of Cairo, uses mostly three-person games to treat the Middle East conflict.

In response to the limits of 2x2 games, some social scientists have passed over extant mathematical techniques and devised systems of their own. Three prominent ideas are Howard's *metagames* (1971), *conflict analysis* (Fraser and Hipel, 1982; Benjamin, 1981), along with *hypergames* (Bennett and Dando, 1970, 1977, 1982), and Brams and Hessels' *theory of moves* (1984) with its *non-myopic equilibria*. (For summaries and comparisons see Hovi, 1987, C. Powell, 1988, Nicholson, 1989, and Zagare, 1986.) These writers share a discontent with the concept of a Nash equilibrium and the narrow assumptions of past theory, but for the wrong reasons, in my view. The conflicts they address are treatable by mainstream game theory. Hypergame analysis addresses situations where one side misunderstands the other's goals, as would incomplete information games. Metagame theory involves adopting general policies of play, as in repeated games. Nonmyopic equilibria apply to players who can switch back and forth in the cells of a matrix, like Danskin's matrix differential games (Washburn, 1977), or some parts of repeated game theory.

The inventors of hypergames and metagames have stressed that their ordinal payoff assumptions are simple enough to apply them to real problems and derive valid decisions. To this end, Howard as well as Fraser and Hipel developed interactive software. Conflict analysis and hypergames have been applied to the Suez crisis (Wright *et al*., 1980; Shupe *et al*., 1980), the fall of France (Bennett and Dando, 1970, 1977), the 1979 Zimbabwe conflict (Kuhn, Hipel and Fraser, 1983), the Cuban

Missile Crisis (Fraser and Hipel, 1982), arms control verification (Cheon and Fraser 1987), and the nuclear arms race (Bennett and Dando, 1982). Some metagame analyses are Richelson's (1979) on nuclear strategy, Alexander's (1976) on Northern Ireland, Benjamin's (1981) on Angola, the Kurdish uprising, and other conflicts, Benjamin's (1982) and Radford's (1988) on the Iranian hostage crisis, and Howard's (1968, 1972) on the Vietnam war, the testban negotiations and the Arab/Israeli conflict (see also Schelling, 1968).

With the theory of moves, Leng (1988) studied crisis bargaining; Brams and Hessel (1984), Solidarity's conflict with the Polish government and the Soviet Union; Zagare (1981, 1983), the Middle East conflicts of 1967 and 1973; Brams (1985), the Cuban missile crisis; and Zagare (1979, 1987a), the Geneva conference of 1954, and the nuclear arms race. Brams (1976) compared a standard and a metagame analyses of the Cuban missile crisis, and Gekker (1989a) did an analysis of the 1948 Berlin crisis using sequential equilibria, and one based on the theory of moves.

Military Operations And Game Theory

Progress in military game theory has gone on mostly unknown to the wider community. While some studies are available in *Operations Research* and *Naval Research Logistics Quarterly*, many have appeared as reports of government laboratories, strategic institutes or private consulting firms. Some of the unclassified American documents can be ordered from the government's National Technical Information Service, but others are simply not publicized, or issued as secret reports, sometimes in secret journals like *The Proceedings of the Military Operations Research Society Symposium* or *The Journal of Defense Research*.
Even game theorists with a clearance can be unaware of their colleagues' work, since secrecy tends to obstruct the circulation of ideas.

Most of the research discussed below is from the United States. I do not know how much has gone on elsewhere, hidden by military secrecy. Some claim that the Soviet Union is more advanced in military game applications, but there seems to be no direct evidence about this either way. Some Soviet texts include military game theory (Ashkenazi, 1961; Suzdal, 1978; Dubin and Suzdal, 1981), but it is in no way deeper than Western works, particularly in regard to its direct usability. In many general books and manuals on military operations research, the game theory is basic but extensive, and is promoted without the warnings so common in the West against applying the results literally (Rehm). By necessity, this review will concentrate on applications in the United States, where access to this information is greatest.

The social value of military game theory has been controversial. To use mathematics to make killing more efficient strikes one as a travesty, and also seems pointless, since each side's gains can usually be imitated by the adversary, with both ending up less secure. On the other hand, the logical analysis might counter governments' natural instinct to build every possible armament in great numbers. The usual argument is basically that for all anyone can say, having the weapon might turn out to confer some benefit some day. Military systems analysis could rationalize planning by prompting officials to think clearly about their goals and objectively evaluate the means. In my view, the most solid policy use of game theory, Kent's Damage Limitation Study, had this benefit. The applications presented here can be judged on their merits.

General discussions with military examples are given by Dresher (1961b), Shubik (1983, 1987a,b), Thomas (1966), and Finn and Kent (1985). The NATO symposium proceedings (Mensch, 1966) give a good sample of problems.

Analyses Of Specific Military Situations

Possibly the earliest attempt on record to solve a real decision problem, beyond board games and gambling, involved World War II anti-submarine patrols. Philip Morse, who had overseen the wartime Anti-submarine Warfare Operations Research Group, posited a sub that must run a corridor, and must travel on the surface for some total time during its passage (Morse and Kimball, 1946). The corridor widens and narrows, making anti-submarine detection less and more effective. Mixed strategies were calculated for how to surface and how to concentrate the search. The measure-theoretic foundations were complex, and it took Blackwell (1954) to add the details. Allentuck (1968) claims that anti-submarine game theory was decisive in winning the Battle of the North Atlantic, but his elaborate evidence is only indirect, and judging from accounts published later (Morse, 1977; Tidman, 1984; Waddington, 1973), the theory probably never directed a real decision.

Probably the most influential analysis of a specific situation was Goldhamer, Marshall and Leites' report on counterforce policy (1959), discussed by O'Neill (1989b). They polled fellow RAND strategists to get estimates of Soviet utilities for various nuclear war outcomes, and described a tree of several stages, without, however, including all the payoffs and probabilities necessary to define the game. The payoff for an American surrender was one of those missing. A more recent model on overall strategic policy is by Phillips and Hayes (1978). A paper interesting for its attention to reality is Hone's matrix game analysis (1984) of the naval arms race during the years between the World Wars. Each side could build ships emphasizing armour, firepower, speed or range, each of these being differentially effective against the other's choice, as in the children's game of Scissors, Paper and Stone. He applied Schelling's technique (1964) of categorizing payoffs as simply high, medium or low. A similar method had been devised independently for a secret 1954 British Army report on the optimum gun/armour balance in a tank (Shephard, 1966).

One episode analysed as a game was later exposed to be a fabrication that was invented to preserve a military secret. The 1944 defeat of the German Seventh Army near Avranche had been attributed to the astute foresight of the American commander, but documents released afterwards showed that the Allies had simply intercepted Hitler's message to his general (Ravid, 1985). The game models of Haywood (1950, 1954) and Brams (1975) would be tidy applications if only the event had really happened.

Game theorists sometimes apologize for recommending mixed strategies, but Beresford and Preston (1955) describe how real combatants came to use them naturally. In daily skirmishes between Malaysian insurgents and British overland convoys, the insurgents could either storm the convoy or snipe at it, and the British could adopt different defensive formations. Each day the British officer in charge concealed a piece of grass in one hand or the other, and a comrade chose a hand, to determine the move. The standard assumption of matrix games, that each acts without knowledge of the other's choice, is questionable in many contexts but appropriate here.

Early Theoretical Emphases

A concern after the Second World War was the theory of *duels*. In the late 1940's at the RAND Corporation, Lloyd Shapley, Richard Bellman, David Blackwell and others introduced important concepts and solutions. Two sides approach each other with bullets of specified accuracies. Each must decide when to fire: shooting early gives you a lower chance of hitting, but waiting too long may mean the opponent shoots first and eliminates you. In alternative versions, the bullets are silent (the opponent does not know when one has been fired), or noisy (the opponent hears your shots, and will know when your ammunition is exhausted). Sometimes fire is continuous, or there are several weapons on each side. The pressing problem of that era was defence against atomic bombers. When should an intercepter open fire on a bomber, and vice versa (Thomas, 1966)? The destructiveness of atomic weapons justified a zero-sum, non-repeated game assumption: the payoff was the probability of destroying the bomber, with no regard to the survival of the fighter. Another application analyzed the optimal detonation height of attacking warheads facing optimally detonating anti-missile interceptors (Kitchen, 1962). Dresher (1961b) contains the basic ideas and Kimmeldorf (1983) reviews recent duelling theory.

Not quite duels, but close to them, were *games of missile launch timing*, investigated at RAND in the late 1950s. The first intercontinental missiles were to be raised out of their vertical protective silos or horizontal `coffins,' prepared and launched. During the preparation time they would be more vulnerable to attack. Several papers (Dresher, 1957; Johnson, 1959; Dresher and Johnson, 1961; see also Brown's camouflaged piece, 1957; Dresher, 1961c, pp. 111-115; and Thomas, 1966) considered the problem of a missile base that chooses a series of firing times, versus an attacker who chooses the times for its warheads to arrive at the base. A warhead destroys a missile with a given probability if it is still in its silo, and with a greater one if it is outside in preparation for firing. The payoff is the number of missiles launched. By 1962 the United States had begun deploying missiles fired directly out of their silos, so the problem lost some relevance.

Games of aiming and evasion (Isaacs, 1954) are still relevant today, but still largely unsolved. A bomber attacks a battleship, which in the simplest version moves back and forth on a line, and turns and accelerates to its maximum velocity instantly. The plane overhead drops a bomb in a chosen direction, and the bomb reaches the one-dimensional ocean surface after a given time lag, destroying anything within a certain distance. The ship knows when the bomb is released but not where it is heading. Finding the approximate mixed strategy solutions is not too difficult: if the radius of destruction is small compared to the interval the ship can reach, the aim point and ship's position should be chosen from a uniform distribution. However a slight change makes the problem more interesting: suppose the ship does not know when the bomb is dropped. It must then make its position continuously unpredictable to the bomber, who is trying to guess it based on observations that stopped a time lag ago (Washburn 1966, 1971). The modern version would involve an intercontinental missile attacking a submarine located by underwater sensors, or a missile-carrying train.

Several papers were written on the strategy of naval mine-laying and detecting. In Scheu's model (1968), the mine-layer uses magnetic influence mines or time-actuated mines left to destroy the searcher, and the searcher may choose to postpone the search and the advance of ships until some of the latter have exploded. Another post-war focus involved *ordnance selection*, such as the armament of a fighter plane that trades off between cannons and machine guns. Fain and Phillips (1964) give an example on the optimal mix of planes on an aircraft carrier. *Reconnaissance* problems received attention (Belzer, 1949; Danskin, 1962a; Dresher, 1957; Sherman, 1949) in regard to the value of acquiring information to be used in a later attack, and the value of blocking the adversary's reconnaissance attempts. Tompkins' *Hotspot game* (Isbell and Marlow, 1955) is an elegant abstract problem suggested by military conflict, involving allocation of resources over time. Each side sends some integral number of resources to a `hotspot' where one or the other loses a unit with a probability determined by the two force sizes. When a unit disappears, either side can send more, until one player's resources are exhausted.

A class of games suggested by warfare is *Colonel Blotto*, involving two players simultaneously dividing their resources among several small engagements. The overall payoff is usually an additive function of who wins each engagement. A simple version was stated by Borel (1921), soon after to assume the post of France's Minister of the Navy. Each player divides one unit among three positions, and whoever has assigned the greater amount to two out of the three positions, wins. Although the game is easy enough to state, some mixed strategy solutions found by Gross and Wagner (1950) have as their supports intricate two-dimensional Cantor sets. Soviet researchers contributed to the theory (Zauberman, 1975), which is reviewed by Washburn (1978). Military applications of Blotto games are fewer than had been expected. The reason seems to be that there are few real contexts where unforeseen numbers of forces from both sides show up at several locations. Exceptions are Shubik and Weber's (1979, 1982) investigation of the defence of a network, in which they introduce payoffs that are not additive in the engagements, and Grotte and Brooks' (1983) discussion of a measure of aircraft carrier presence. Another application of Blotto-like games is the problem of missile attack and defence, to be treated next.

Missile Attack And Defence

Missile defence theory is not only mathematically advanced, it had a far-reaching influence on policy. A prominent model is *Prim-Read theory*, named for its two originators. In a simple version, an attacker sends missile warheads to destroy a group of fixed targets, and a defender tries to protect them using interceptors, which are themselves missiles. At the first stage the defender divides its interceptors among sites, and plans how many interceptors to send up from each targeted site against each incoming warhead. The attacker, knowing the allocation and the firing plans, divides its warheads among the sites, and launches them sequentially at their targets. If an attacking warhead gets through, it destroys the site with certainty, but each interceptor succeeds only with a probability. There can be no reallocation: if the attacker has destroyed a site, further weapons that had been assigned there cannot be sent elsewhere, nor can interceptors be shifted. The payoff usually involves the sum of the values of the targets destroyed.

The recommended rule for allocating the defence interceptors, the Prim-Read deployment, typically does not divide the defences in proportion to the values of the targets, but if we allow fractional allocations, it induces the attacker to send warheads in those proportions. A remarkable result is that defensive deployments exist that are best no matter how many weapons the attacker possesses, so the defender need not know the opposing force size. The basics were laid out by Read (1957, 1961), and Karr (1981) derives some uniqueness and optimality results. Examples of modern applications are the papers of Holmes (1982) and Bracken and Brooks (1983).

Prim-Read theory was crucial in a report that may have helped dampen the arms competition. In 1964, US Air Force General Glenn Kent completed his secret `Damage-Limitation Study' on how to make nuclear war less destructive. One prospect was building a vast anti-missile system. With Prim-Read theory, Kent's group derived a supportable estimate of a defence's effectiveness. The verdict was negative: an anti-missile system would be expensive, and could be negated at lower marginal cost to the adversary by increasing the attacking force. Kent's study promoted the case for a ban on large anti-missile systems (Kaplan, 1983), and in 1972 the two powers signed the Anti-Ballistic Missile Treaty. It continues in force today, and represents probably the greatest success of arms control, a move to a `Pareto-superior equilibrium' at which the two sides avoid a more intense race in defensive and offensive weapons.

While few policy decisions on record were influenced by game theory, we have to separate innate reasons for this from institutional ones. Perhaps the opportunity was there to apply the theory, but the right person was not in a place to carry it through. Kent possessed a reputation for innovativeness, had associated with Schelling at Harvard, and written on the mathematics of conflict (1963). When he came into the position of influence, he used the theory effectively.

The most questionable assumption of Prim-Read theory for the Damage Limitation Study was that the attacker knows the defender's firing schedule, but the opposite premise would require a mixed strategy solution. Mixed strategies would have been more complicated, and were foreign to the operations researchers who addressed these problems, who were more accustomed to optimization. McGarvey (1987), however, recently investigated this case. Another branch of missile defence theory alters another part of Prim-Read, by assuming that the attacker is unaware of the defender's allocation. The attitude has been that Prim-Read theory treats attacks against population centres, and the alternative branch deals with attacks on missile sites. Strauch's models (1965, 1967) assumed perfectly functioning interceptors, and thus were formally Blotto games. Other versions (Matheson, 1966, 1967) had imperfect interception, so each side had a strategic decision beyond the Blotto allocation, of how many interceptors to send up against each attacking warhead, not knowing how many more would be descending on the particular target. Each side calculates the best `taper,' decreasing numbers of weapons to be assigned to the various targets, and randomly matches the numbers with the targets. Unlike the Prim-Read case, a deployment that is ideal against an all-out attack may be poor against a smaller one, and Bracken, *et al*. (1987) have recently found `robust' defence allocations that for any attack level, stay within a certain percentage of the optimum. Burr *et al*. (1985) find solutions without the approximation of continuously divisible missiles. Matlin (1970) reviewed missile defence models, about a dozen of which involved games, and Eckler and Burr (1972) wrote a thorough report on research up to that time. Many of the studies they described should have treated their situations as games, but used one-person optimization, and so were forced to assume inflexible or irrational behaviour by one player.

Ronald Reagan's vision of anti-missile satellites shielding an entire country bypassed the mathematically fertile study of allocation among defended sites. New game models appeared more slowly than one would expect from the bountiful funds available, but recent examples are by O'Meara and Soland (1988), as well as several authors they reference, on the coordination problem of guarding a large area. Another theme is the timing of the launch of heat decoys by an offense facing a boost-phase anti-missile system.

Tactical Air War Models

Strategy for air operations has preoccupied many military operations researchers, and for good reason. In a European war, each side would be continually dividing its fighter-bombers between attacking the adversary's ground forces, intercepting its aircraft and bombing its airfields. The two sides' decisions would be interdependent, of course, since one's best choice today depends on what it expects the other to do in the future. The aircraft allocation rule might influence the course of a conventional war in Europe, but analysts of the `conventional balance' (e.g., Posen, 1984; Epstein, 1987) have ignored it. Some large computer models stress air activity, some the ground war, often as a function of the sponsoring military service, but few address the best way to set their interaction. An exception was the programme TAC CONTENDER, forerunner of the current RAND programme TAC SAGE, which attempted to produce an optimal strategy, but Falk (1973) demonstrated a counterexample.

Satisfactory tactical air models might bolster each bloc's feeling of security, and its willingness to negotiate force reductions. However the problem has so far been solved only in simplified versions. Research began quite early (Giamboni, Mengel and Dishington, 1951), and was stimulated by Berkovitz and Dresher's success in solving (1959, 1960; Dresher, 1961b) a complicated sequential game of perfect information. The problem is to make a realistic problem computationally manageable, and there are three approaches: Lagrange multipliers as in TAC CONTENDER, which provide only a sufficient condition for optimality, grid methods, applied by George Dantzig (Control Analysis Corporation, 1972), with the same shortcoming, and solutions of trees of matrix games (Bracken, Falk and Karr, 1975; Schwartz, 1979), which are sure to be optimal but are feasible only for games of fewer stages.

Lanchester Models Of Fire Control

Lanchester's much-used equations of combat assume a homogeneous force on each side. Extending them to allow a mixed force, like infantry and artillery, requires an assumption about how each component will be allocated against the adversary's components. Almost always the approach used in practice is to divide the forces in some arbitrary proportion, which remains fixed throughout the battle, but this policy might be beatable if the opponent concentrates and shifts the target of each component in a clever way. Accordingly, one stream of game models treats the allocation-of-fire decision as a differential game constrained by Lanchester's equations. In Weiss's initiating work (1957, 1959), each side could direct its artillery against the other's artillery or ground forces, but the ground forces attacked only each other. Whereas airwar applications are usually discrete, with planes sent out daily, fire control games are continuous. Taylor (1974) made Weiss's results more rigorous, and Kawara (1973), followed up by Taylor (1977, 1978) and Taylor and Brown (1978), adds an interesting variant in which the attacker's ground forces shoot while they move towards the defender's line, but with `area fire,' not precisely aimed because they are moving. The artillery bombardment must cease when the two sides meet, so that each side tries for the highest ratio of its forces to the other's at that time. Moglewer and Payne (1970) treat fire control and resupply jointly (see also Sternberg, 1971, and Isaacs' Game of Attack and Attrition, 1965), and the field is summarized by Taylor (1983, Ch.8).

Search And Ambush

In a *search game*, a Searcher tries to find a Hider within some time limit, or in the shortest time. Sometimes the Hider is moving, like a submarine, sometimes it is fixed, like a mine. Even though the research began in World War II, it has found few applicable results, perhaps because of the difficulty of incorporating two or three-dimensional space. The bulk of applied search theory today is only `one-sided', postulating a non-intelligent evader located in different places with exogenous probabilities. Several dozen game-oriented papers, solving abstract problems, are listed by Dobbie (1966) and the Chudnovskys (1988). Gal's book (1980) analyses many theoretical games. One variant requires the evader to call at a port within a time interval (Washburn, 1971), and Dobbie (1966) suggested a game where both sides are searchers, ships trying to rendezvous with their convoy.

Related to search games are continuous *games of ambush*, investigated by Ruckle in his book (1983) and in a series of articles in *Operations Research*. Usually the Ambusher does not move, so the game is easier to solve than a search. A simple example has the Evader choosing a path from one side of the unit square to the other, while the Ambusher selects a set of prescribed area in the square, and receives as payoff the length of the path that lies in the area. A less elegant but more practical example (Randall, 1966) involves anti-submarine barriers. Danskin's paper on convoy routing (1962b) is also relevant.

Pursuit Games

Pursuit problems (Isaacs, 1965; Hayek, 1980) were studied in the 1950s to improve defence against atomic bombers, and in fact sparked the whole area of differential games. Pursuit game theory has grown steadily, and a recent bibliography (Rodin, 1988) lists hundreds of articles. Typically the Pursuer tries to come within capture distance of the Evader, but in some games, the angle of motion is important, as in the case of manoeuvring fighters. Practical pursuit problems are hard to solve and hard to apply, since human behaviour cannot be programmed second-by-second. Aerial dogfights involving complete turns are too complicated, and add the wrinkle that the Evader is trying to switch roles with the Pursuer. However, strategies optimal for sections of an engagement have been calculated, with the goal of abstracting general rules (Ho, 1970). A more successful application is the control of manoeuvring nuclear warheads and interceptor missiles. The short engagement times and the interceptors' high accelerations mean less manoeuvring, so their movements can be represented by simpler kinematic laws. Despite their amenability to mathematical analysis, the only manoeuvring warheads currently facing anti-missile interceptors are British missiles aimed at Moscow. Discussions of differential games in military operations research are given by Isaacs (1975), Ho and Olsder (1983) and Schirlitzki (1976).

Values Of Weapons

An issue of practical interest is how to assign numbers to represent the military worth of weapons. The goal is to generate indices of overall military strength, in order to assess the `military balance,' to direct strategic decisions in terms of what resources should be `traded' for what, to score war games, or to recommend a compromise in arms control agreements. Some procedures work from the bottom up, estimating the qualities of individual weapons based on their design features, then adding to evaluate the whole arsenal, but clearly additivity is a questionable assumption here, since some weapons complement each other, others are substitutes. A game-theoretical approach might determine the benefits of having an arsenal and infer back to the worth of the components. Pugh and Mayberry (1973 -- see also Pugh, 1973) treat the question of the proper objective function of war, following Nash's general bargaining model, to argue that each side will try for the most favourable negotiated settlement by conducting war as strictly competitive. O'Neill (1988a) applies the Shapley value, regarding the weapons as players in a coalitional game. The characteristic function displays non-monotonicity whenever the enemy's weapons join the coalition. These models do not lead to a usable measure, but do clarify the informal debate.

Command, Control, And Communication

Some writers have analyzed the contest of a transmitter and receiver versus a jammer. Fain's early report (1961) on the tradeoff between offensive forces and jamming units was panned by his fellow conference participants, but other studies followed it. Some writers discuss a jammer who chooses a signal with a limit on its bandwidth and power (Weiss and Schwartz, 1985; Stark, 1982; Helin, 1983; Basar, 1983; Basar and Wu, 1985; Bansal and Basar, 1989.) Others posit a network where the transmitter can route the communication a certain way, and the jammer attacks certain nodes or links (Polydoros and Cheng, 1987). Further game analyses look at the authentication of messages (Simmons, 1981; Brickell, 1984). The papers above may be the tip of a larger classified literature.

McGuire (1958) assumed unreliable communication links connecting ICBMs and discussed the tradeoff between retaliating when one should not, e.g., in response to an accidental breakdown of a link, versus not retaliating after a real attack. For a given network, the attacker plans what to target and the defender plans what instructions for retaliation to issue to base commanders who find themselves incommunicado. Independently, Weiss (1983) considered missiles that retaliate against targets of differing values, where each launch control officer may retarget based on partial knowledge of what other missiles have been destroyed. The retaliator issues plans for the missile commander, contingent on the latter's knowledge of what others remain, and the striker decides what to attack. In Weiss' view, implementing the theory would bolster a government's deterrent threat without adding weapons that it could use in a first strike. In other words, the proposal would avoid the security dilemma. Notwithstanding, the Strategic Air Command has generally not transferred this kind of control down to the level of a missile base commander.

Sequential Models Of A Nuclear War

Several models of a nuclear war involve the attacker dividing its weapons between the adversary's population and retaliating missiles. The sides maximize some objective functions of the damage to each after a fixed number of attacks. Usually the solution method is a max-min techniques, but one exception is Dalkey's study (1965). Early papers used Lagrange techniques on two-stage models and so guaranteed only local optimality, but Bracken, Falk and Miercort (1977) found a method for the globally optimal solution in a class of reasonable problems. Grotte's paper (1982) gives an example with many references. One complex model by Bracken (1989a) has twelve stages, where each side allocates missiles among the roles of attacking population targets, attacking military targets and saving as reserve forces, and also decides how much of its missile defence system to hold back for use against future strikes. A discrete game of perfect information with 20,000,000 possible paths, it is surely the most complicated tree yet solved by brute force. The solution takes about three minutes on a Cray computer.

An attempt was made to incorporate two-sided optimal choice in the Arsenal Exchange Model, the computer programme most used by the United States government to investigate nuclear war strategies, but the task turned out to be too complex, and it is still the user who chooses levels of attack and priorities of targets, whereupon a linear programming routine assigns weapons to targets.

Military And IR Applications Compared

Most IR game modellers began as social scientists, and acquired mathematics on their own. Relative to those few who were trained in mathematics and turned to the social sciences later, their models tend to reflect a more worldly intuition and greater interest in problems important for policy. They have been less liable to apply formal methods prematurely, since they have usually been discussing the matter at length with their mainstream colleagues. On the other hand, many of the models surveyed here are primitive, sometimes with an incorrect solution, or none at all. Many are at the level of `proto-game theory' (O'Neill, 1989b), using the concepts but no formal derivations. The field has yielded many single models but no edifice of theorems.

The low technical state of IR game theory so far can be blamed partly on its subject matter. There is more to a game model than filling in a matrix -- many of the better ones spot some simple regularity in the situation and state it as a mathematical assumption. However, many IR situations lack obvious `mathematical handles.' Unlike economics, for example, the conflict is not directly over goods or money. These commodities make modelling easier by giving a structure to players' preferences, through assumptions of continuity or risk aversion, or conservation of the sum of the commodities when they are divided among the players. After many economics transactions the parties simply walk away, but an international game may have no clear endpoint. It is exactly because of these obstacles that as IR game theory becomes more conceptually sophisticated, it will become more interesting. The mathematics will have to be different than economics or social choice applications, and new theoretical concepts will arise. The challenge will be to keep its results understandable to the non-mathematical majority.

Military applications have been more advanced mathematically than those in IR, perhaps because they deal with physical rather than social entities: objects move around in space, destroy each other with rules set by the physics of weaponry, or forces are measurable and can be divided up among missions. Also, compared to international politics, the goal is clearer: to seize territory, or to maximize some function of adversary forces destroyed and one's own preserved. Military analysts can often construe the game as zero-sum, and thereby derive a unique solution.

In regard to many IR papers, it seems odd to develop a theory without deriving its predictions, but one often encounters IR games that are stated but not solved. This situation is rarer in military game theory, which is less concerned with clarifying concepts, more with choosing moves. However no military model can include enough factors to be precisely true, so analysts have sought principles of good strategy (Dresher, 1966), general patterns of behaviour recommended by the solution, to use in situations close to the one analyzed. Their approach resonates with the idea of military doctrine. An example of such a principle is the target-defence principle of Prim-Read theory, that one should never assign defences so as to induce the attacker to assault a defended target while sending none against some undefended one. This idea is related to the `no soft-spot principle' in general studies of defence allocation. Another example comes from tactical airwar models, which continually suggest that the superior force should split its attack, while the inferior should concentrate on a point randomly chosen.

Prospects For IR Game Theory

The reference section includes about twenty publications with IR game models from the 1960s, fifty from the 1970s and 180 from the 1980s. These data reflect the activity of practitioners, but general enthusiasm has not followed a rising curve. In the early days many saw game theory as the science of strategy appropriate to the atomic era, when wars must be planned deductively, rather than based on historical cases. Many were opposed to this method, which apparently requires one to formulate goals explicitly, and prescribes what to do in complex situations without a role for experiences, intuitions or attitudes if they cannot be formalized. Its mathematical character suggested that the same technocrats that gave the world the nuclear arms race, were now seeking control over decisions on peace and war. The name `game theory' proved an unfortunate choice, suggesting that war was a frivolous pastime where one side was trying simply to beat the other, and confusing the discipline with the non-mathematical field of military gaming. Some statements are those of Maccoby (1961), Aron (1962), and Green (1966) as attackers; McDonald (1949), Wohlstetter (1964) and Kaplan (1973) as defenders; and Schelling (1960) and Rapoport (1964, 1965, 1966) as reformers. Oddly enough, concurrent with this deep suspicion that game theory was a tool of the Cold War, many in the peace research community became enthusiastic about it.

Interest waned at the end of the 1960s, but game models left one important truth impressed on many minds: that international struggles are not zero-sum, that intense competition still allows cooperative acts. From the late 1970s, the books and articles of Brams and Kilgour helped move the method back into view. The 1980s saw a surge of papers, and increasing interest especially among recent graduates, but the attitude of the majority of political scientists is equivocal. The staying influence of Schelling's writings have given game theory a presumption of importance, and Axelrod's work has promoted access to it by the non-mathematical majority. In security studies, game applications offer abstract reasoning, which, it is broadly agreed, the field needs. On the other hand, there is no mass movement to learn the necessary mathematics, and graduate students who wish to learn it, often have nowhere to go. Thousands of political science undergraduates hear about PD each year, but the set of real practitioners remains small. Many IR scholars misunderstand it profoundly. They reject games as the ultimate rational actor theory, explain at length its `innate limitations', or, like Blackett quoted above, castigate it because it cannot make precise numerical predictions, a test no other theoretical approach could pass.

How should we judge Blackett's point that if game theory were practical, practical people would be using it? His premise is that any theory inapplicable in a simple context like gambling, is *ipso facto* irrelevant to more complicated international issues. In fact game theory clarifies problems of peace and war just because they are more complicated. Unlike card games, even the basic concepts of discussion are obscure, even the rules are uncertain. What does it mean to `show resolve'? What constitutes `escalation'? How can cooperation emerge from international anarchy? Especially in IR, the contribution of game models will be to sort out these conceptual problems.

References